Version 1.0

Metadata Requirements – Base Reference for NASA Earth Science Data Products

1 Status of this Document

This document provides information to the NASA Earth Science Division (ESD). This document specifies an ESD implementation for NASA Earth Science metadata. Distribution of this document is unlimited.

2 Change Explanation

N/A

3 Abstract

This document defines the NASA Earth Science Division base metadata requirements for science data products. The metadata requirements represented in this document are a means of assuring the consistency of data requirements across subsystems, and supporting the data standardization necessary for system interoperability. The International Organization for Standardization (ISO) Geographic Information – Metadata standard 19115 (and related standards) shall be used to describe science data products. This standard includes two parts, 19115 and 19115-2. Part 2 and revisions to 19115 that are currently being developed by ISO, include elements that are critical for NASA collections and products. In addition, the data quality elements of 19115 have been extracted into another standard (ISO 19157). Complete descriptions of NASA collections and products will require elements from all of these standards.

4 Table of Contents

5 Introduction

Metadata serves a multitude of purposes. Missions typically list information about the instrumentation, the flight vehicle, the data gridding method, the temporal and spatial extent of the data, the operational plan and flight dynamics information. Data Centers employ metadata to enable product ordering, subsetting, merging and visualization. In addition, metadata provide users with descriptions of data product structure, lineage, algorithms and the quality of the content.

The role of metadata elements vary based on their scope. The following terms are frequently used to categorize metadata by scope:

> **Series/Collection metadata** – Metadata elements that describe an entire set of data products or files. Values of series/collection metadata apply to all of the products in a specific aggregate. Series/collection metadata may represent the same release of any given data product, sets of data generated during an experiment, a campaign or an algorithm test.
>
> **Dataset/Granule metadata** – Metadata elements that describe a single instance of a data product. Values of dataset/granule metadata apply to all of the data in that one product. Typical metadata in this category describe spatial and temporal extent of the data instance as well as the quality and lineage.

The use of term "dataset" in NASA ESDS metadata conflicts with ISO terminology. In ISO, "dataset" means "granule". In NASA ESDS, "dataset" means "collection". Series/collection and dataset/granule metadata are used in search and discovery of data products. A metadata repository, commonly referred to as a catalog, is accessed by client applications that support user queries. A small set of metadata attributes is sufficient for the most common spatial, temporal and parameter based queries.

ESD science data products shall contain metadata conforming to the ISO 19115 Geographic Information – Metadata standard. This standard includes two parts, 19115 and 19115-2. Part 2 and revisions to 19115 that are currently being developed by ISO, include elements that are critical for NASA collections and products. In addition, the data quality elements of 19115 have been extracted into another standard (ISO 19157). Complete descriptions of NASA collections and products will require elements from all of these standards. The XML encoding of ISO 19115 metadata is specified in ISO 19139, which shall also be considered part of these requirements.

6 Base Reference Metadata Model

The ESD base metadata requirements outline the minimum metadata elements required to adhere to NASA Satellite Mission Data Systems Requirements.

The flexibility of ISO 19115 enables assignment of metadata to multiple entities and elements in multiple configurations. Usage of a common set of metadata entities and elements enhances interoperability across NASA missions. The entities and elements presented in this document represent a subset of what is required for NASA science data products.

NASA specific ISO 19115 conventions shall be posted to the EOSDIS website along with tools, guides, documentation and example implementations of the ESD metadata requirements.

(http://earthdata.nasa.gov/)

As the targets of metadata description the following three constructs are utilized in this document. Although individual mission/instruments may utilize different terms for these construct, the corresponding descriptions should all be addressed.

- **Series/Collection** - A grouping of science data that all come from the same source, such as a modeling group or institution. Series/collections have information that is common across all the datasets/granules they contain.
- **Dataset/Granule** - The smallest aggregation of data that can be independently managed (described, inventoried, and retrieved). Datasets/granules have their own metadata model and support values associated with the additional attributes defined by series/collections that they are part of.
- **Browse** - An image that provides a high-level view of the associated dataset/granule or series/collection metadata item. Browse do not have an independent representation or data model, but are contained within the Series/Collection or Dataset/Granule metadata.

A series/collection may contain zero or more datasets/granules, however datasets/granules cannot exist without being associated with a series/collection. Browse images may be associated or included within either series/collections or datasets/granules.

For each metadata type, the minimum metadata entities or elements required to describe science data products are outlined below. The metadata name and description of each metadata entity or element are listed in tabular form. In instances where top-level entities of complex metadata elements are listed, the requisite child element information is inherently required to correctly represent the parent metadata element.

The metadata names displayed in this document are not required to match the metadata names used in the mission/project data products. However, the data products shall provide elements that include the required content.

It is recognized that there are mission/project dependent elements that are not included in this document but should be treated as required elements for the mission/project.

SERIES / COLLECTIONS

Name	Description	Standard
MD_DataIdentification.citation/CI_Citation.title	This element specifies a name for the series/collection.	19115:2003
MI_Metadata.fileIdentifier	This element specifies a unique identifier for the series/collection.	19115:2003
MD_DataIdentification.citation/CI_Citation.edition	This element specifies the version identifier of the data series/collection	19115:2003
MD_DataIdentification.abstract	This element identifies the purpose and provides a description of the content of the series/collection.	19115:2003
MD_DataIdentification.purpose	This element provides a summary of the intentions with which the data were developed	19115:2003
EX_Extent	This entity describes the spatial and temporal extent of the data	19115:2003
MD_Distribution	This entity describes data format, transfer options and size.	19115:2003
SV_ServiceIdentification	This entity describes access, subsetting or processing services available for the dataset/granule as well as operations supported by those services and parameters for controlling their behavior.	19115:2003
MI_AcquisitionInformation	This entity describes the platforms associated with the acquisition of the series/collection or dataset/granule, instrument/sensor used to measure or record data and acquisition plan.	19115:2003
MI_CoverageDescription	This entity describes the physical parameters being measured or calculated (bands).	19115:2003
MD_Keywords	This element allows for the specification of Earth science keywords that are representative of the series/collection. The Science Keyword list is managed by the Global Master Change Directory (GCMD).	19115:2003
MI_Operation	This entity contains attributes describing the scientific endeavor(s) to which the series/collection is associated. Scientific endeavors include campaigns, projects, interdisciplinary science investigations, missions, field experiments, etc.	19115:2003
CI_OnlineResource	This entity describes online resources that supplement the information included in the structured metadata.	19115:2003
CI_ResponsibleParty	This element contains elements for describing people or organizations that are related to the data and their roles.	19115:2003
MD_BrowseGraphic	List of browse images associated with the series/collection.	19115:2003
MD_DataIdentification.citation	This entity allows data users to properly cite the series/collection.	19115:2003

Name	Description	Standard
LI_Lineage	This entity records the data product generation in detail sufficient to allow reproducibility.	19115:2003
DQ_Element	This entity records the information about the quality of the data or any quality assurance procedures followed in producing the data.	19115:2003
MD_Constraints	This entity provides information pertaining to any constraints for access.	19115:2003
MD_GridSpatialRepresentation	This entity describes the spatial representation.	19115:2003

Entities contain multiple elements and the referenced standard must be consulted to identify the required elements.

DATASETS / GRANULES

Datasets/granules shall incorporate all fields listed in the series/collections table in addition to the elements listed below.

Name	Description	Standard
MI_Metadata.fileIdentifier	The Universal Reference (UR) ID of the dataset/granule. This ID is unique to the data product.	19115:2003
DS_Series	This element holds information about series/collections that a dataset/granule is part of.	19115:2003

7 Authors' Address

NASA Earth Science Data and Information System (ESDIS), Goddard Space Flight Center (GSFC) Code 423.0 Greenbelt Road, Greenbelt Maryland 20771 USA